ASHA American Speech-Language-Hearing Association

JSLHR

Research Note

Physiological Correlates of the Spectral Aggregate of the High-Passed Fundamental Frequency in Adductor Laryngeal Dystonia

Laura E. Toles,[a] Katherine L. Marks,[b,c] Manuel E. Díaz Cádiz,[b] Taylor F. Feaster,[b] Andrew Wrede,[b] Maanya Baranwal,[b] Anastasia Haubrich,[b] Misora Ito,[b] Sophie Li,[b] Bonnie Little,[b] Samantha Lonergan,[b] Franklin Phan,[b] Giavanna Siracusano,[b] Alex Estrada,[b] Jose M. Rojas,[b] Daniel P. Buckley,[b,d] Tanya Eadie,[e] Cara L. Sauder,[e] Andrew Keltz,[f] Michael Lerner,[f] Pavan Mallur,[g] Gregory Grillone,[d] J. Pieter Noordzij,[d] Lauren F. Tracy,[d] Ted Mau,[a,h] Cara E. Stepp,[b,d,i] and Jenny Vojtech[b]

[a] Department of Otolaryngology - Head and Neck Surgery, The University of Texas Southwestern Medical Center, Dallas [b] Department of Speech, Language and Hearing Sciences, Boston University, MA [c] Department of Otolaryngology, Emory University School of Medicine, Atlanta, GA [d] Department of Otolaryngology–Head & Neck Surgery, Boston University School of Medicine, MA [e] Department of Speech and Hearing Sciences, University of Washington, Seattle [f] Department of Otolaryngology, Yale School of Medicine, Greenwich, CT [g] Department of Otolaryngology–Head and Neck Surgery, Harvard Medical School, Boston, MA [h] Department of Otolaryngology—Head & Neck Surgery, Stanford University School of Medicine, Palo Alto, CA [i] Department of Biomedical Engineering, Boston University, MA

ARTICLE INFO

Article History:
Received October 15, 2025
Revision received December 22, 2025
Accepted January 26, 2026

Editor-in-Chief: Jessica E. Huber
Editor: Supraja Anand

https://doi.org/10.1044/2026_JSLHR-25-00836

ABSTRACT

Objective: The spectral aggregate of the high-passed fundamental frequency (SAHf_o) is an automated acoustic measure associated with the primary acoustic discontinuities present in the speech of individuals with adductor laryngeal dystonia (AdLD). Acoustic discontinuities (e.g., phonation breaks, frequency shifts, creak) are also associated with supraglottic compression and visual laryngeal obstruction postures that occur during laryngeal spasms (i.e., phasic or tonic laryngeal dystonic activity). The purpose of this study was to determine whether SAHf_o is similarly reflective of these physiological manifestations of AdLD.
Method: Thirty-two speakers with AdLD read voiced and voiceless phoneme-loaded sentences while undergoing transnasal laryngoscopy. Simultaneous microphone and high-speed videoendoscopy signals were recorded. SAHf_o was automatically calculated from the acoustic signal, and laryngeal postures (visual laryngeal obstructions, anterior–posterior and mediolateral supraglottic compression) were manually labeled. A generalized linear mixed-effects analysis assessed the effects of sentence type (voiced, voiceless), obstruction events (obstructed view of vocal folds, visible vocal folds), mediolateral compression, anterior–posterior compression, and the interaction of Mediolateral × Anterior–Posterior compression on SAHf_o.
Results: Sentence type, obstruction events, and mediolateral compression were significantly related to SAHf_o. Anterior–posterior compression did not have a significant effect, but there was a significant interaction between mediolateral and anterior–posterior compression. SAHf_o was significantly lower when both mediolateral and anterior–posterior compression were mild and significantly higher when mediolateral compression was severe, regardless of presence or severity of anterior–posterior compression.
Conclusion: Findings suggest that SAHf_o is related to the laryngeal postures associated with AdLD-specific acoustic discontinuities, providing preliminary physiological support of the use of SAHf_o in speakers with AdLD.

Correspondence to Laura E. Toles: laura.toles@utsouthwestern.edu.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Involuntary spasms of the adductory laryngeal muscles and a strained-strangled voice quality during speech are hallmark signs of adductor laryngeal dystonia (AdLD; Mor & Blitzer, 2016). These signs are associated with distinct

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

acoustic discontinuities such as phonatory breaks, frequency shifts, and creak (i.e., low or irregular fundamental frequency [f_o]; Brown et al., 2000; Marks et al., 2022, 2023, 2024; Sapienza et al., 1998). Perceptual and acoustic expression of AdLD is heterogeneous, and some individuals exhibit a predominantly strained-strangled voice quality with few or no overt phonatory breaks, consistent with a more tonic manifestation of the disorder. In such cases, abnormal voice presentation may arise from sustained or tonic hyperadduction that produces acoustic irregularities without discrete interruptions of voicing. For the purposes of this research note, the term *laryngeal spasm* is used broadly to refer to laryngeal dystonic activity in AdLD, encompassing both phasic spasmodic events and more sustained or fluctuating hyperadductory postures that contribute to a strained-strangled voice quality.

Individuals with AdLD often report high levels of vocal effort during speech (Baylor et al., 2005), and laryngeal spasms can trigger compensatory vocal hyperfunction patterns such as supraglottic compression and aberrant laryngeal postures (Chen et al., 2020; Leonard & Kendall, 1999; Woodson et al., 1991). Therefore, individuals with AdLD might also be diagnosed with secondary muscle tension dysphonia. Vocal tremor is also a frequent concomitant diagnosis, occurring in 26% to 60% of cases (Hintze et al., 2017; White et al., 2011). Clinical diagnosis of AdLD involves a combination of auditory-perceptual judgments, acoustic assessment, patient self-report scales, and laryngeal endoscopy. Unfortunately, traditional auditory-perceptual, acoustic, and laryngoscopic evaluations are not always sensitive to the primary signs of AdLD (i.e., laryngeal spasms, strained-strangled voice quality, and associated discontinuities) as AdLD can be misdiagnosed as other disorders involving vocal hyperfunction (i.e., muscle tension dysphonia; R. Patel et al., 2011). A primary limitation to reliably diagnosing AdLD and measuring treatment response is the lack of robust, physiologically validated objective measures that are sensitive to the primary signs and characteristics of the disorder. Without the means for clinically feasible, objective measurement of signs specific to AdLD, it is difficult to make appropriate diagnoses and measure response to treatment.

The gold standard for identifying acoustic discontinuities associated with AdLD is the manual labeling of discontinuities in the acoustic signal by a trained technician (Sapienza et al., 1998, 1999, 2000, 2002). This method is extremely time intensive and therefore not clinically feasible. As a surrogate for manual labeling, Buckley et al. (2020) proposed an outcome measure later termed the spectral aggregate of the high-passed fundamental frequency (SAHf_o; Marks et al., 2022) as an automated objective measure with promise to correlate with AdLD-specific acoustic discontinuities. SAHf_o is automatically calculated by detection of fast transitions in f_o in voiced segments (i.e., $\geq$ 50 Hz within 50 ms) and discontinuities in f_o caused by voice breaks or removal of unvoiced segments. In the initial study (Buckley et al., 2020), a combination of SAHf_o and three relative fundamental frequency (RFF) measures were included in a statistical model that found that the combined measures represented 85% of the variance in auditory-perceptual judgments of overall severity of dysphonia. Since RFF measures are associated with laryngeal tension and vocal effort (McKenna et al., 2016; McKenna & Stepp, 2018), the authors postulated that the RFF measures explained the variance in the overall severity of dysphonia related to compensatory hyperfunctional behaviors. SAHf_o independently had a strong positive correlation (r = .70) with overall severity of dysphonia, suggesting greater SAHf_o values could correspond with more severe auditory-perceptual presentation of the primary symptoms of AdLD: phonatory breaks and discontinuities. In the follow-up study by Marks et al. (2022), increased SAHf_o was significantly related to manually identified AdLD-specific acoustic discontinuities presumably caused by laryngeal spasms, which provided concurrent validity for the measure. Of note, in that same study, greater SAHf_o values occurred with voiceless phoneme-loaded sentences when compared to voiced phoneme-loaded sentences. This phenomenon likely occurred because SAHf_o calculations remove voiceless sounds, and thus, voiceless phoneme-loaded content has more opportunity for abrupt changes in f_o to occur. As such, it is important that studies using SAHf_o should control for phonemic context.

Although SAHf_o appears to be related to acoustic discontinuities, it has yet to be physiologically validated as being associated with laryngeal postures that correspond with laryngeal spasms. A recent study by Marks et al. (2025) found that manually labeled acoustic discontinuities were statistically related to the presence of supraglottic compression and visual laryngeal obstructions of the vocal folds identified with high-speed videoendoscopy (HSV). That study represented the first step in identifying noninvasive outcome measures in the acoustic signal that are potentially related to laryngeal spasms. Since manual labeling of acoustic discontinuities is cumbersome and time consuming, the next logical step is to determine whether SAHf_o is associated with laryngeal postures consistent with adductory spasms in AdLD.

There is a substantial need to validate clinically feasible, automatic acoustic measures that are associated with the physiological underpinnings of AdLD. The objective of this study is to determine the concurrent validity of SAHf_o as an outcome measure that reflects the physiological characteristics of AdLD. Considering that SAHf_o has been found to correlate with the primary acoustic signs of AdLD and that those acoustic signs are associated with

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

laryngeal obstructions and supraglottic compression common in AdLD, we hypothesized that within an AdLD cohort, SAHf_o would be positively associated with visual laryngeal obstructive postures and supraglottic compression patterns identified through HSV.

Materials and Method

Participants

Thirty-two speakers (21 cisgender females, 11 cisgender males) participated in this study. All speakers had been diagnosed with AdLD by a board-certified, fellowship-trained laryngologist. Diagnoses were made based on consensus criteria from Ludlow et al. (2018), including intermittent glottal stops or vowel breaks on voiced sentences; strained-strangled, effortful, tight voice quality; patient report of speaking effort; symptoms reduced during whisper; typical structure and symmetry of the vocal folds at rest on laryngoscopy; and/or intermittent vocal fold or arytenoid hyperadduction on laryngoscopy. Speakers with concomitant vocal tremor were retained for analysis as long as they had a documented diagnosis of AdLD.[1] Exclusion criteria included diagnosis of other voice disorders (e.g., primary muscle tension dysphonia, structural lesions). Speakers' ages ranged from 32 to 87 years (females: 68.0 ± 12.0 years; males: 57.1 ± 14.3). They exhibited a range of AdLD severity and were receiving ongoing Botox treatment for their disorder. To ensure that speakers were symptomatic at the time of testing, all recordings took place shortly before their scheduled Botox injection when participants confirmed they were symptomatic. Participants provided informed consent to voluntarily participate in accordance with the Boston University (BU) Institutional Review Board (Protocol #2625) or The University of Texas Southwestern (UTSW) Medical Center Institutional Review Board (Protocol #STU-2023-0965). Fourteen speakers participated at BU and 18 at UTSW.

Three voice-specialized speech-language pathologists blindly rated the overall severity of dysphonia of all 32 speakers using the 100-mm visual analog scale from the Consensus Auditory-Perceptual Evaluation of Voice with nonlinearly placed textual severity labels (American Speech-Language-Hearing Association, 2002). The mean overall severity of dysphonia was 46.2 (*SD* = 14.9, range: 16.6–86.7). Interrater reliability using the intraclass correlation coefficient (ICC) for a two-way mixed-effects model (ICC(3,k)) was .82, 95% CI [0.79, 0.94].

[1]A voice-specializing speech-language pathologist who was blinded to the study conditions evaluated whether or not each participant presented with co-occurring vocal tremor using a combined sample of sustained vowel production, a voice-loaded sentence, and a voiceless-loaded sentence (Johnson et al., 2014). Based on these ratings, 13 of 32 participants (41%) had co-occurring tremor, which is consistent with previous reports in the literature (A. B. Patel et al., 2015; Wolraich et al., 2010).

Procedure

Participants underwent simultaneous acoustic and HSV recordings in a sound-treated booth. Participants seen at BU received aerosolized nasal decongestant without topical anesthetic, and those participating at UTSW received aerosolized nasal decongestant with topical lidocaine approximately 5–10 min prior to endoscopy. An ear-mounted directional microphone was set at mouth-to-microphone distance of approximately 7 cm off-center from the lips. A flexible laryngeal endoscope (BU: PENTAX Medical, Model FNL-7RP3, 2.4 mm; UTSW: Olympus Model ENF-P4, 3.6 mm) was attached to a high-speed camera (BU: FASTCAM Mini AX100I, Model 540 K-C-16GB; UTSW: Phantom V311 [Vision Research]) and passed through the nasal cavity to obtain laryngoscopic images. Images were captured at 1,000 frames per second at BU and 500 frames per second at UTSW with 256 x 256 pixel resolution and a steady xenon light source (BU:300-W KayPentax Model 7162B; UTSW: 400-W Titan 400E). As described in our previous work (Marks et al., 2025), all participants read aloud 14 unfamiliar complex sentences, displayed in Table 1. Sentences were constructed to elicit laryngeal spasms using syntactic and linguistic complexity while also allowing for analysis of RFF in future projects. Eight of the sentences were loaded with primarily voiced phonemes and the other six with primarily voiceless phonemes.

Acoustic Analysis

Consistent with previous studies (Buckley et al., 2020; Marks et al., 2022), automated SAHf_o measures were calculated using Halcyon—an instantaneous f_o estimator—and custom scripts in MATLAB 23.2 (MathWorks). Automated processing involved aligning audio recordings to stimuli text using the Penn Phonetics Lab Forced Aligner (Yuan & Liberman, 2008) and subsequent removal of unvoiced sounds. When speakers misspoke or substituted words (e.g., saying "on which" instead of "where"), words were transcribed as actually produced rather than as they appeared in the target sentence. An f_o contour was estimated by optimizing the most probable path across a range of f_o candidates, set to a range of 30- to 400-Hz over 1-ms sampling windows. The resulting contour was resampled using linear interpolation to match the timescale of the original acoustic signal. Then, spectrograms were calculated using the f_o contour as a function of time, with a 25-ms Hanning window and 90% overlap. After calculating the time-average mean from the spectrogram, information was high-pass filtered by aggregating energy above 1000 Hz. For the time interval corresponding to each phoneme, the accumulated

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 1. Sentence stimuli read aloud by participants during high-speed videoendoscopy.

Voiced phoneme-loaded sentences	Voiceless phoneme-loaded sentences
The leafy tree where the dew shimmered was where she bravely married the man.	A shiny key with a true sheen opens a spiffy shoe case in her home.
The mushy banana that Annie feeds Nan amuses Macy every day.	The cheeky kid kicked his squeaky chew toy, so Sheila put the sticky tape to the tip of it.
My good buddy Goofy that used eighty shiny new shoe bins delivered wine one day.	Should she see Kasey put her feet on the sofa tonight, Fifi can tell Pete to pay a hefty fee for his sin.
The beefy meat that the wee teen enjoyed was with a goofy new fan of New England.	We feasted two cheeky teens sushi and tofu stew when she came home Saturday.
We ate messy eggs in New Orleans on greasy oatmeal in their shiny metal bowls.	Sixty-two soups found their place at her table with two fishy sides at her sushi cart.
The girl Lisa that Mandy meets in the busy scene dreamed of owning blue shimmery nails again.	Tracy will pay for Lucy to fetch treats from the seaside with Stacey and Pete.
When we gave Vicky a key to the shack, the knife on the beam was a major deal.	
Murphy sees Mia near the new bin in Lisa's big, blue sailboat at dawn on Monday morning.	

energy was normalized by the number of frequency bins, resulting in one estimate of SAHf_o per phoneme.

HSV Analysis

Using the same methods from our recent study (Marks et al., 2025), two trained technicians used a custom script in MATLAB 23.2 for manual identification of laryngeal postures of visual obstruction of the vocal folds and supraglottic compression in the HSV exam recordings for each of the 14 sentences. HSV labels were made at the phoneme level. Visual obstruction labels were applied to phonemes if greater than two thirds of the vocal folds were covered by the arytenoid cartilages or the epiglottis. Phonemes with supraglottic compression were classified based on the direction (anterior–posterior [A-P] or mediolateral) and degree of compression based on how much of the vocal folds were obstructed by the surrounding tissue (< 1/3 was mild, 1/3–2/3 was moderate, > 2/3 was severe). The two technicians convened to review and reach consensus on any labels for which they disagreed. Of note, two voiceless-loaded sentences could not be analyzed for one participant due to sustained epiglottic obstruction preventing adequate visualization throughout those sentences. An additional two sentences from a separate participant were excluded due to technical errors during data processing. In total, four out of 448 sentences (0.9%) were excluded from analysis. All remaining sentences across all 32 participants were successfully analyzed. This phoneme-level approach yielded 15,300 voiced phonemes for subsequent analysis, nested within 444 sentences across 32 participants.

Statistical Analyses

All statistical analyses were carried out in Minitab (v.20.4). A generalized linear mixed-effects analysis was conducted to assess the effects of sentence type (voiced, unvoiced), obstruction events (dummy coded; 1 = obstructed view of the vocal folds, 0 = visible vocal folds), mediolateral compression (none, mild, moderate, severe), A-P compression (none, mild, moderate, severe), and the interaction of Mediolateral × A-P compression on SAHf_o. Since there are multiple planes of supraglottic compression (i.e., mediolateral and A-P), we included their interaction to determine how the interplay between severity and direction of supraglottic compression influenced SAHf_o. Speaker and phoneme were included as random effects to account for inherent individual differences and differences within phonemic context. Significance was set a priori to $p < .05$, and partial eta squared (η_p^2) values were calculated as effect sizes. Post hoc Dunnett tests were performed for significant main effects to examine multiple comparisons between fixed effect levels.

Results

The residuals of the initial linear mixed-effects model violated assumptions of normality. A Box–Cox transformation was therefore applied to the dependent variable (SAHf_o) to normalize residuals and stabilize variance. All main effects demonstrated significant effects on the transformed SAHf_o data, with the exception of A-P compression, $F(3, 15248) = 0.35$, $p = .723$ (see Table 2). SAHf_o was significantly higher in sentences loaded with voiceless phonemes compared to those loaded with voiced phonemes, $F(1, 15242) = 35.47$, $t = -5.96$, $p = .001$, and when the vocal folds were obstructed compared to when they were visible, $F(1, 15247) = 38.54$, $t = 6.21$, $p < .001$.

Figure 1 shows the estimated marginal means of the transformed SAHf_o data for significant main effects, and

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

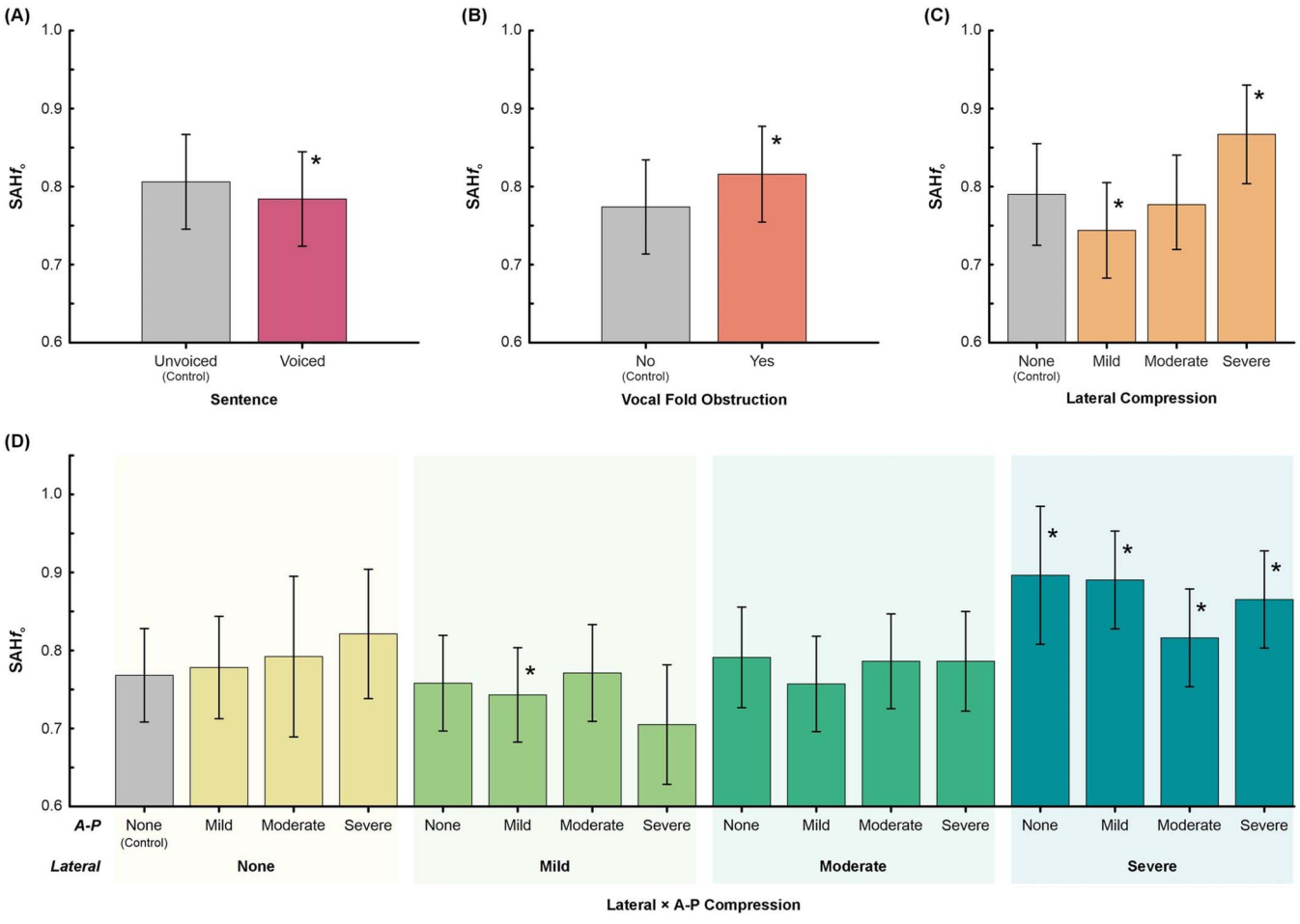
Table 2. Linear mixed-effects model results on spectral aggregate of the high-passed fundamental frequency.

Effect	*df*	*F*	*p*	η_p^2
Sentence type	1, 15242	35.47	< .001	.002
Obstruction event	1, 15247	38.54	< .001	.003
Mediolateral compression	3, 15239	31.57	< .001	.006
A-P compression	3, 15248	0.35	.788	< .001
Lateral × A-P Compression	9, 15236	4.51	< .001	.003

Note. A-P = anterior–posterior.

Table 3 presents descriptive statistics for the transformed and back-transformed SAHf_o data to aid interpretation. Trends in SAHf_o varied in relation to mediolateral compression, $F(3, 15239) = 31.57$, $p < .001$. Specifically, SAHf_o estimates were significantly greater when mediolateral compression was severe ($t = 4.46$, $p < .001$) but significantly smaller when mild ($t = -3.04$, $p = .006$). A similar pattern emerged in the interaction between mediolateral and A-P compression, $F(9, 15236) = 4.51$, $p < .001$. SAHf_o remained significantly greater whenever mediolateral compression was severe, regardless of A-P compression status (all $p < .05$). Additionally, SAHf_o values

Figure 1. Estimated marginal means for spectral aggregate of the high-passed fundamental frequency (SAHf_o) based on (A) sentence type, (B) vocal fold obstruction, (C) lateral (i.e., mediolateral) compression, and (D) Mediolateral × A-P Compression. Asterisks (*) indicate statistical significance ($p < .05$) relative to the control level, shown in gray and labeled "(Control)." Error bars represent 95% confidence intervals. A-P = anterior–posterior.



Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 3. Estimated marginal means for Box–Cox transformed and back-transformed spectral aggregate of the high-passed fundamental frequency values.

Variable & levels	Transformed values *M* [95% CI]		Back-transformed values *M* [95% CI]	
Sentence type				
Unvoiced (control)	0.806	[0.745, 0.867]	0.141	[0.045, 0.237]
Voiced	0.784	[0.724, 0.844]	0.109	[0.033, 0.186]
Obstruction event				
No (control)	0.774	[0.714, 0.834]	0.097	[0.029, 0.166]
Yes	0.816	[0.755, 0.877]	0.157	[0.050, 0.265]
Mediolateral compression				
None (control)	0.790	[0.725, 0.855]	0.117	[0.029, 0.205]
Mild	0.744	[0.683, 0.805]	0.068	[0.017, 0.119]
Moderate	0.780	[0.720, 0.840]	0.104	[0.031, 0.178]
Severe	0.867	[0.804, 0.930]	0.273	[0.092, 0.454]
A-P compression				
None (control)	0.803	[0.741, 0.865]	0.136	[0.040, 0.232]
Mild	0.792	[0.731, 0.853]	0.120	[0.036, 0.204]
Moderate	0.791	[0.727, 0.855]	0.119	[0.032, 0.206]
Severe	0.794	[0.731, 0.857]	0.123	[0.034, 0.212]
Mediolateral × A-P Compression				
None × None (control)	0.768	[0.708, 0.828]	0.091	[0.027, 0.155]
None × Mild	0.778	[0.713, 0.843]	0.102	[0.024, 0.180]
None × Moderate	0.792	[0.689, 0.895]	0.120	[–0.021, 0.262]
None × Severe	0.821	[0.738, 0.904]	0.166	[0.014, 0.319]
Mild × None	0.758	[0.697, 0.819]	0.081	[0.021, 0.140]
Mild × Mild	0.743	[0.683, 0.803]	0.067	[0.018, 0.117]
Mild × Moderate	0.771	[0.709, 0.833]	0.094	[0.025, 0.163]
Mild × Severe	0.705	[0.629, 0.781]	0.042	[0.001, 0.083]
Moderate × None	0.791	[0.727, 0.855]	0.119	[0.031, 0.206]
Moderate × Mild	0.757	[0.696, 0.818]	0.080	[0.021, 0.138]
Moderate × Moderate	0.786	[0.725, 0.847]	0.112	[0.034, 0.190]
Moderate × Severe	0.786	[0.722, 0.850]	0.112	[0.029, 0.195]
Severe × None	0.896	[0.808, 0.984]	0.368	[0.039, 0.698]
Severe × Mild	0.890	[0.827, 0.953]	0.347	[0.125, 0.568]
Severe × Moderate	0.816	[0.753, 0.879]	0.157	[0.048, 0.267]
Severe × Severe	0.865	[0.803, 0.927]	0.268	[0.093, 0.442]

Note. CI = confidence interval; A-P = anterior–posterior.

were significantly lower when both mediolateral and A-P compression were mild ($t = -3.57$, $p = .005$).

Discussion

The goal of this study was to provide further evidence of the concurrent validity of an automated acoustic outcome measure, SAHf_o, for use in the AdLD population. Specifically, since laryngeal spasms are a pathogenetic characteristic associated with AdLD, we aimed to investigate whether SAHf_o is related to the physiological laryngeal obstruction postures and supraglottic compression events that co-occur with laryngeal spasms during connected speech (Leonard & Kendall, 1999). We hypothesized that SAHf_o would increase during visual obstruction and compression postures. The findings support that hypothesis as both visual obstruction and mediolateral supraglottic compression were associated with greater SAHf_o levels.

Previous studies on SAHf_o have provided evidence supporting its use as an outcome measure related both to perceived overall severity of dysphonia (Buckley et al., 2020) in individuals with AdLD as well as its association with AdLD-specific acoustic discontinuities (Marks et al., 2022). A recent study by our group also found that visual vocal fold obstructions and supraglottic compression were significantly related to manually labeled acoustic discontinuities (Marks et al., 2025), which provided preliminary insights into the physiological underpinnings of acoustic

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

discontinuities during running speech. The current study builds off these previous studies. Results show that SAHf_o is associated with the supraglottic manifestations of laryngeal spasms. Although laryngeal spasms are often conceptualized as discrete events, AdLD is increasingly understood as a disorder of atypical laryngeal motor control, with both phasic and tonic manifestations contributing to voice symptoms. As such, SAHf_o can be used as an outcome measure to represent the presence and severity of laryngeal obstruction and mediolateral compression postures during dystonic laryngeal activity in AdLD.

Although the findings support the use of SAHf_o as a physiologically based outcome measure for AdLD, it is important to note that SAHf_o performs differently on voiced phoneme-loaded stimuli compared to voiceless phoneme-loaded stimuli (Marks et al., 2022). The current results highlight this phenomenon, as sentence type was a significant main effect. Voiceless phoneme-loaded sentences tended to have greater levels of SAHf_o, which is consistent with the previous study investigating SAHf_o (Marks et al., 2025). This result might seem counterintuitive since laryngeal spasms tend to occur more frequently during voiced phonemes in AdLD. In the automated algorithm, SAHf_o excludes voiceless sounds when identifying changes in the f_o contour. Despite more laryngeal spasms occurring during voiced phoneme-loaded stimuli, the f_o contour naturally changes more abruptly around voiceless phonemes, thereby leading to greater SAHf_o levels. As such, it is important to consider the phonemic content of speech stimuli when conducting SAHf_o analysis. For instance, it would not be valid to compare SAHf_o at baseline using voiced phoneme-loaded stimuli with posttreatment SAHf_o values calculated using voiceless phoneme-loaded stimuli. Such a comparison could erroneously indicate that no treatment effect occurred. Hence, comparisons should only be made across stimuli with similar or identical phonemic content.

Degree and direction of supraglottic compression is often considered a sign of compensation for vocal hyperfunction or other laryngeal anomalies (Behrman et al., 2003; Bielamowicz et al., 2004). A study by Leonard and Kendall (1999) found that mediolateral supraglottic compression occurred frequently in both AdLD and muscle tension dysphonia. In that study, compression occurrences were more episodic during continuous speech in AdLD, occurring more markedly on all voiced segments. That result provided evidence to suggest that supraglottic compression co-occurs with phonatory breaks and related discontinuities in AdLD. Our recent study found that manually labeled acoustic discontinuities were significantly related to supraglottic compression (Marks et al., 2025). Interestingly, that work did not find a statistically significant interaction between group (i.e., AdLD vs. healthy control) and supraglottic compression. Direction of supraglottic compression (i.e., A-P and mediolateral) was not differentiated in that study. Literature has shown that supraglottic compression is present in both vocally healthy and voice disorder populations (Stager et al., 2000), but the severity and/or direction of compression might differ. For instance, A-P compression in individuals with muscle tension dysphonia has been comparable to vocally healthy controls, but some studies indicate that mediolateral compression can be more severe in muscle tension dysphonia (Shembel et al., 2025; Toles & Shembel, 2024). These past findings suggest that A-P compression is not necessarily specific to pathological presentations, whereas more severe levels of mediolateral compression could be more indicative of an aberrant laryngeal posture, particularly if it results in other atypical presentations (e.g., changes in voice quality, vocal effort). Similarly, results from the current study found that A-P compression was not significantly related to SAHf_o. However, severe levels of mediolateral compression were significantly related to higher SAHf_o measurements, regardless of the level of A-P compression. Therefore, our findings suggest that severe mediolateral compression is associated with acoustic discontinuities in AdLD. This further supports the notion that mediolateral compression might be more pronounced during pathological phonation, particularly that which occurs concurrently with vocal fold hyperadduction. However, this claim will need to be further supported by future work that employs a control group and speakers with muscle tension dysphonia for comparison.

Limitations and Future Directions

The current study used lower HSV sampling rates than other studies using HSV (i.e., 500 and 1,000 frames per second versus 4,000+ frames per second; R. Patel et al., 2011). Although our sampling rates were not sufficient to capture rapid features and micromotions of the mucosal wave during vibration, changes in laryngeal postures occur on a timescale of 104–227 ms (Dailey et al., 2005). Thus, the sampling rate is adequate for the purposes of the current study. Future investigations that are interested in the relationship between SAHf_o and vocal fold vibratory kinematic measurement should incorporate higher HSV sampling rates.

Participants at one site (UTSW) received topical lidocaine in addition to nasal decongestant prior to endoscopy, whereas participants at the other site (BU) received only nasal decongestant. Topical anesthesia has been reported to potentially alter sensorimotor feedback and serve as a sensory trick (geste antagoniste) in some individuals with AdLD (Dwenger et al., 2025), which could influence symptom expression during the recording. While our statistical model accounted for speaker-level variability through random effects, we cannot rule out potential site-related differences in symptom presentation due to this procedural variation.

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

SAHf_o is calculated based on detection of rapid f_o transitions and discontinuities caused by voice breaks. Therefore, it may be less sensitive to predominantly tonic manifestations of AdLD characterized by sustained hyperadduction and strained-strangled voice quality without overt phonatory breaks or rapid f_o shifts. Future work should examine whether SAHf_o patterns differ across the spectrum of AdLD presentations and explore complementary measures that may capture tonic hyperadductory postures.

This study did not include speakers without voice disorders or individuals with non-AdLD voice disorders. As such, the findings should be interpreted as reflecting within-group associations in individuals with AdLD. Given that supraglottic compression has been documented in healthy voices and in other voice disorders and that epiglottic obstruction may occur in typical speech for articulatory purposes, we interpret SAHf_o as a measure of the severity of physiological features within individuals with AdLD rather than as a marker unique to this population. Future studies incorporating control groups will be needed to determine whether the observed physiological correlates of SAHf_o differ across populations. While this study established phoneme-level physiological validation of SAHf_o, future research should examine speaker-level relationships between overall dysphonia severity, SAHf_o, and supraglottic compression to inform diagnostic thresholds and severity-stratified treatment outcomes.

Conclusions

This study provided preliminary physiological validation of SAHf_o as an outcome measure in AdLD. Findings revealed that SAHf_o was related to the presence of visual laryngeal obstructions and mediolateral supraglottic compression, both common laryngeal expressions in AdLD (Leonard & Kendall, 1999; Marks et al., 2025; Woodson et al., 1991). Visual laryngeal obstructions and mediolateral compression were each associated with higher SAHf_o measurements. SAHf_o values were highest when mediolateral compression was severe, regardless of presence or severity of A-P compression. A-P compression was unrelated to SAHf_o. Considering manually labeled AdLD-specific acoustic discontinuities are related to SAHf_o, visual laryngeal obstructions, and supraglottic compression, the current study supports the use of SAHf_o as an automated acoustic outcome measure reflecting both acoustic and physiologic signs of AdLD. Further research should investigate the use of SAHf_o to determine its sensitivity and responsiveness to change for further validation of its use as an outcome measure in the AdLD population.

Data Availability Statement

The data sets analyzed during the current study are not publicly available due to identifiable information.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders (DC015570, awarded to Cara E. Stepp, and DC020349, awarded to Katherine L. Marks), the Boston University Sargent College Student Research Grant (awarded to Katherine L. Marks), and the American Speech-Language-Hearing Foundation Speech Science Grant (awarded to Katherine L. Marks). The authors acknowledge the lab managers, research fellows, graduate students, and research assistants who assisted in recruitment, screening, and scheduling of participants and supported laboratory visits.

References

American Speech-Language-Hearing Association. (2002). *Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)*. Retrieved January 13, 2026, from https://www.asha.org/form/cape-v/ [PDF]

Baylor, C. R., Yorkston, K. M., & Eadie, T. L. (2005). The consequences of spasmodic dysphonia on communication-related quality of life: A qualitative study of the insider's experiences. *Journal of Communication Disorders, 38*(5), 395–419. https://doi.org/10.1016/j.jcomdis.2005.03.003

Behrman, A., Dahl, L. D., Abramson, A. L., & Schutte, H. K. (2003). Anterior-posterior and medial compression of the supraglottis: Signs of nonorganic dysphonia or normal postures? *Journal of Voice, 17*(3), 403–410. https://doi.org/10.1067/S0892-1997(03)00018-3

Bielamowicz, S., Kapoor, R., Schwartz, J., & Stager, S. V. (2004). Relationship among glottal area, static supraglottic compression, and laryngeal function studies in unilateral vocal fold paresis and paralysis. *Journal of Voice, 18*(1), 138–145. https://doi.org/10.1016/j.jvoice.2003.11.005

Brown, W. S., Jr., Rothman, H. B., & Sapienza, C. M. (2000). Perceptual and acoustic study of professionally trained versus untrained voices. *Journal of Voice, 14*(3), 301–309. https://doi.org/10.1016/S0892-1997(00)80076-4

Buckley, D. P., Cadiz, M. D., Eadie, T. L., & Stepp, C. E. (2020). Acoustic model of perceived overall severity of dysphonia in adductor-type laryngeal dystonia. *Journal of Speech, Language, and Hearing Research, 63*(8), 2713–2722. https://doi.org/10.1044/2020_JSLHR-19-00354

Chen, W., Woo, P., & Murry, T. (2020). Vibratory onset of adductor spasmodic dysphonia and muscle tension dysphonia: A high-speed video study. *Journal of Voice, 34*(4), 598–603. https://doi.org/10.1016/j.jvoice.2018.12.010

Dailey, S. H., Kobler, J. B., Hillman, R. E., Tangrom, K., Thananart, E., Mauri, M., & Zeitels, S. M. (2005). Endoscopic measurement of vocal fold movement during adduction and abduction. *The Laryngoscope, 115*(1), 178–183. https://doi.org/10.1097/01.mlg.0000150701.46377.df

Dwenger, K., Roy, N., Jennings, S. G., Smith, M. E., Mathy, P., Simonyan, K., & Barkmeier-Kraemer, J. M. (2025). Comparing the effects of sensory tricks on voice symptoms in patients with laryngeal dystonia and essential vocal tremor. *Journal of Speech, Language, and Hearing Research, 68*(4), 1654–1675. https://doi.org/10.1044/2024_JSLHR-24-00476

Hintze, J. M., Ludlow, C. L., Bansberg, S. F., Adler, C. H., & Lott, D. G. (2017). Spasmodic dysphonia: A review. Part 1:

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Pathogenic factors. *Otolaryngology—Head and Neck Surgery, 157*(4), 551–557. https://doi.org/10.1177/0194599817728521

Johnson, D. M., Hapner, E. R., Klein, A. M., Pethan, M., & Johns, M. M., III. (2014). Validation of a telephone screening tool for spasmodic dysphonia and vocal fold tremor. *Journal of Voice, 28*(6), 711–715. https://doi.org/10.1016/j.jvoice.2014.03.009

Leonard, R., & Kendall, K. (1999). Differentiation of spasmodic and psychogenic dysphonias with phonoscopic evaluation. *The Laryngoscope, 109*(2), 295–300. https://doi.org/10.1097/00005537-199902000-00022

Ludlow, C. L., Domangue, R., Sharma, D., Jinnah, H. A., Perlmutter, J. S., Berke, G., Sapienza, C., Smith, M. E., Blumin, J. H., Kalata, C. E., Blindauer, K., Johns, M., Hapner, E., Harmon, A., Paniello, R., Adler, C. H., Crujido, L., Lott, D. G., Bansberg, S. F., . . . Stebbins, G. (2018). *Consensus-based attributes for identifying patients with spasmodic dysphonia and other voice disorders. JAMA Otolaryngology—Head & Neck Surgery, 144*(8), 657–665. https://doi.org/10.1001/jamaoto.2018.0644

Marks, K. L., Díaz Cádiz, M. E., Toles, L. E., Buckley, D. P., Tracy, L. F., Noordzji, J. P., Grillone, G. A., & Stepp, C. E. (2023). Automated creak differentiates adductor laryngeal dystonia and muscle tension dysphonia. *The Laryngoscope, 133*(10), 2687–2694. https://doi.org/10.1002/lary.30588

Marks, K. L., Feaster, T. F., Baker, S., Díaz-Cádiz, M. E., Doyle, P. C., & Stepp, C. E. (2022). Spectral aggregate of the high-passed fundamental frequency and its relationship to the primary acoustic features of adductor laryngeal dystonia. *Journal of Speech, Language, and Hearing Research, 65*(11), 4085–4095. https://doi.org/10.1044/2022_JSLHR-22-00157

Marks, K. L., Frankford, S. A., Cocroft, S. J., Lonergan, S., Díaz Cádiz, M. E., & Stepp, C. E. (2024). The prevalence of creak across breath groups in adductor laryngeal dystonia. *Journal of Voice*. Advance online publication. https://doi.org/10.1016/j.jvoice.2024.10.013

Marks, K. L., Vojtech, J., Díaz Cádiz, M. E., Toles, L. E., Feaster, T. F., Wrede, A., Baranwal, M., Haubrich, A., Ito, M., Li, S., Little, B., Lonergan, S., Phan, F., Siracusano, G., Estrada, A., Rojas, J. M., Buckley, D. P., Sauder, C. L., Keltz, A., . . . Stepp, C. E. (2026). Observation of laryngeal postures via high-speed videoendoscopy in adductor laryngeal dystonia. *Annals of Otology, Rhinology & Laryngology, 135*(4), 268–279. https://doi.org/10.1177/00034894251388829

McKenna, V. S., Heller Murray, E. S., Lien, Y.-A. S., & Stepp, C. E. (2016). The relationship between relative fundamental frequency and a kinematic estimate of laryngeal stiffness in healthy adults. *Journal of Speech, Language, and Hearing Research, 59*(6), 1283–1294. https://doi.org/10.1044/2016_JSLHR-S-15-0406

McKenna, V. S., & Stepp, C. E. (2018). The relationship between acoustical and perceptual measures of vocal effort. *The Journal of the Acoustical Society of America, 144*(3), 1643–1658. https://doi.org/10.1121/1.5055234

Mor, N., & Blitzer, A. (2016). *Diagnosis and treatment of laryngeal dystonia: Past, present and future directions*. Tremor and Other Hyperkinetic Movements, 6. https://doi.org/10.7916/D8G160J5

Patel, A. B., Bansberg, S. F., Adler, C. H., Lott, D. G., & Crujido, L. (2015). The Mayo Clinic Arizona spasmodic dysphonia experience: A demographic analysis of 718 patients. *Annals of Otology, Rhinology & Laryngology, 124*(11), 859–863. https://doi.org/10.1177/000348941558855

Patel, R., Liu, L., Galatsanos, N., & Bless, D. M. (2011). Differential vibratory characteristics of adductor spasmodic dysphonia and muscle tension dysphonia on high-speed digital imaging. *Annals of Otology, Rhinology & Laryngology, 120*(1), 21–32. https://doi.org/10.1177/000348941112000104

Sapienza, C. M., Cannito, M. P., Murry, T., Branski, R., & Woodson, G. (2002). Acoustic variations in reading produced by speakers with spasmodic dysphonia pre-botox injection and within early stages of post-botox injection. *Journal of Speech, Language, and Hearing Research, 45*(5), 830–843. https://doi.org/10.1044/1092-4388(2002/067)

Sapienza, C. M., Murry, T., & Brown, W. S., Jr. (1998). Variations in adductor spasmodic dysphonia: Acoustic evidence. *Journal of Voice, 12*(2), 214–222. https://doi.org/10.1016/S0892-1997(98)80041-6

Sapienza, C. M., Walton, S., & Murry, T. (1999). Acoustic variations in adductor spasmodic dysphonia as a function of speech task. *Journal of Speech, Language, and Hearing Research, 42*(1), 127–140. https://doi.org/10.1044/jslhr.4201.127

Sapienza, C. M., Walton, S., & Murry, T. (2000). Adductor spasmodic dysphonia and muscular tension dysphonia: Acoustic analysis of sustained phonation and reading. *Journal of Voice, 14*(4), 502–520. https://doi.org/10.1016/S0892-1997(00)80008-9

Shembel, A. C., Morrison, R. A., McDowell, S., Smeltzer, J. C., Crocker, C., Bell, C., & Mau, T. (2025). Relationships between laryngoscopic analysis metrics of supraglottic compression and vocal effort in primary muscle tension dysphonia. *Journal of Voice, 39*(6), 1571–1578. https://doi.org/10.1016/j.jvoice.2023.06.011

Stager, S. V., Bielamowicz, S. A., Regnell, J. R., Gupta, A., & Barkmeier, J. M. (2000). Supraglottic activity: Evidence of vocal hyperfunction or laryngeal articulation? *Journal of Speech, Language, and Hearing Research, 43*(1), 229–238. https://doi.org/10.1044/jslhr.4301.229

Toles, L. E., & Shembel, A. C. (2024). Acoustic and physiologic correlates of vocal effort in individuals with and without primary muscle tension dysphonia. *American Journal of Speech-Language Pathology, 33*(1), 237–247. https://doi.org/10.1044/2023_AJSLP-23-00159

White, L. J., Klein, A. M., Hapner, E. R., Delgaudio, J. M., Hanfelt, J. J., Jinnah, H. A., & Johns, M. M., III. (2011). Coprevalence of tremor with spasmodic dysphonia: A case-control study. *The Laryngoscope, 121*(8), 1752–1755. https://doi.org/10.1002/lary.21872

Wolraich, D., Vasile Marchis-Crisan, C., Redding, N., Khella, S. L., & Mirza, N. (2010). Laryngeal tremor: Co-occurrence with other movement disorders. *ORL, 72*(5), 291–294. https://doi.org/10.1159/000317032

Woodson, G. E., Zwirner, P., Murry, T., & Swenson, M. (1991). Use of flexible fiberoptic laryngoscopy to assess patients with spasmodic dysphonia. *Journal of Voice, 5*(1), 85–91. https://doi.org/10.1016/S0892-1997(05)80168-7

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America, 123*(5), Article 3878. https://doi.org/10.1121/1.2935783

Downloaded from: https://pubs.asha.org Boston University on 07/10/2026, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions